Metal-insulator transition in 2D: resistance in the critical region

B. L. Altshuler[1,2], D. L. Maslov,[3] and V. M. Pudalov[4]

[1] *NEC Research Institute, 4 Independence Way, Princeton, NJ 08540*
[2] *Physics Department, Princeton University, Princeton, NJ 08544*
[3] *Department of Physics, University of Florida*
P. O. Box 118440, Gainesville, Florida 32611-8440
[4] *P. N. Lebedev Physics Institute, Russian Academy of Sciences*
Leninsky Prospect 53, Moscow 117924, Russia
(February 1, 2008)

The goal of this paper is to highlight several issues which are most crucial for the understanding of the "metal-insulator transition" in two dimensions. We discuss some common problems in interpreting experimental results on high mobility Si MOSFETs. We analyze concepts and methods used to determine the critical density of electrons at the metal-insulator transition. In particular, we discuss the origin of the temperature dependence of the resistivity and reasons for this dependence to flatten out at some electron density in the vicinity of the metal-insulator transition. This flattening has recently been proposed to indicate a true quantum phase transition. We suggest an alternative interpretation of this result and demonstrate the consistency of our proposition with the experimental data. One of the main questions, which arise in connection with the transition, is whether or not the metallic state is qualitatively distinct from a conventional disordered Fermi liquid. We analyze the arguments in favor of both affirmative and negative answers to this question and conclude that the experimental results accumulated up-to-date do not provide convincing evidence for the new state of matter characterized by a metallic-like residual conductivity. We also discuss in details the measurement and control of the *electron* temperature; these issues are crucial for interpreting the low-temperature experimental data.

PACS numbers: 71.30.+h, 72.15 Rn, 73.40.Qv

I. INTRODUCTION

After several years of intensive experimental and theoretical efforts (see, e.g., Ref. [1] for an extensive bibliography), even the basic features of the phenomenon known as the "metal-insulator transition in two dimensions"(2D MIT) remain to be the subjects of ongoing discussions. Is this phenomenon a true "quantum phase transition"(QPT) [2] or can it be understood in terms of conventional physics of disordered conductors [3,4] ? This question is at the heart of the whole discussion. Recently, we wrote a paper [5] in favor of the second possibility. In particular, we argued that

i) in the "metallic phase", 2D systems seem to behave as a quite conventional disordered metal rather than a distinctly new state of matter and

ii) it is possible to explain the anomalous behavior of the resistivity, $\rho(T)$, by an interplay of two temperature dependences: the one given by a metallic-like, quasiclassical (Drude) resistivity [6], $\rho_{\rm d}(T)$, and the other one arising from quantum localization.

We do realize that ii) represents a rather naive approach, at least because it does not fully take into account the Coulomb interaction between electrons, whereas the most pronounced effects have been observed in systems with *a priori* strong electron-electron correlations. Nevertheless, this approach allows one to describe qualitatively, and even semi-quantitatively, most of the results on electron transport both in metallic phase and near the transition point. This suggests that it is a good starting point for developing an adequate theoretical understanding of the MIT in two dimensions.

Conclusion i) is based on the analysis of existing experimental data, as summarized in Refs. [5,7]. It has been further supported by recent experimental papers [1,8–10], which showed that the transport properties of Si/Ge, Si MOSFET and p-GaAs structures in the metallic regime can be successfully interpreted in conventional terms. Authors of Refs. [9,10] have also identified the contribution of electron-electron interactions to the resistance (via measuring the temperature-dependent part of the Hall resistance) and found that it remains quite small even when parameter r_s is rather large. There is also a number of recent theoretical papers [11–14,7] which, although differing in a particular mechanism for the T-dependence of the Drude resistivity, share the general spirit of propositions i) and ii).

An alternative point of view, expressed, for example, in Refs. [15–17], is that the observed MIT-like behavior indicates a true quantum phase transition between an insulator and a novel metallic phase. Motivated by the importance of this controversy, i.e., QPT versus propositions i) and ii), for the field of quantum transport in 2D and its possible relevance for other realizations of quantum phase transitions, we decided to analyze in detail the arguments on both sides of the QPT question. In this paper, we discuss recent and some of the previously published experimental results on the resistivity of Si MOS-FETs in the vicinity of the MIT, along with some of the theoretical interpretations of these results. We conclude that that there is no convincing experimental evidence for the QPT in the existing data.

As one of the main questions in the field is the behavior of a 2D system in the limit $T \to 0$, we begin our analysis with a short summary of recent results by Kravchenko and Klapwijk (KK) [15], who measured the temperature dependence of the resistivity down to a bath temperature as low as $T_{\rm bath} = 35\,{\rm mK}$. This paper reports measurements on a single Si MOSFET sample with peak mobility $\mu_{\rm peak} = 27,000\,{\rm cm}^2/{\rm Vs}$ in the range from $35\,{\rm mK}$ to $1.2\,{\rm K}$ at five different electron concentrations. Digitized data from Ref. [15] is shown in Fig. 1 (curves *1-5*).

In Ref. [15], the following two points are emphasized:

1) For $n = n_1 = 6.85 \times 10^{10}{\rm cm}^{-2}$ and $n = n_2 = 7.17 \times 10^{10}{\rm cm}^{-2}$ (curves *1* and *2*), the resistivity decreases as temperature increases ($d\rho/dT < 0$); for $n = n_4 = 7.57 \times 10^{10}{\rm cm}^{-2}$ and $n = n_5 = 7.85 \times 10^{10}{\rm cm}^{-2}$ (curves *4* and *5*), the resistivity increases with temperature ($d\rho/dT > 0$). Based on these observations, the authors argue that curves *1* and *2* correspond to an insulating phase, whereas curves *4* and *5* demonstrate a metallic behavior. At $n = n_3 = 7.25 \times 10^{10}{\rm cm}^{-2}$ (curve *3*), only a weak ($\pm 5\%$) temperature dependence was observed within the interval of bath temperature from $35\,{\rm mK}$ to $1\,{\rm K}$. The authors conclude that $n_3 = 7.25 \times 10^{10}{\rm cm}^{-2}$ is the critical electron density for their sample, i.e., it corresponds to the metal-insulator transition point. Note that this conclusion is based entirely on the temperature independence of curve *3* over a relatively narrow temperature range.

2) Another observation, emphasized in Ref. [15], is that in the range $0.1\,{\rm K} < T < 0.4\,{\rm K}$ the metallic $\rho(T)$ - dependences (curves *4, 5*) appear to be nearly linear and thus different from the exponential temperature dependence observed in other experiments. KK stress that in this region the resistivity shows no insulating up-turn at the lowest temperature achieved.

KK consider the fact that they observed no temperature dependence of the resistivity of their sample at a certain concentration $n = n_3$ as a strong argument against our model [5]. They point out that such a precise cancellation would require a special relation between $\rho_d(T)$ and the scaling $\beta(\rho)$-function. KK argue that this relation between two objects of entirely different origin "would be a remarkable coincidence", and is therefore unlikely.

Based on the arguments listed above, authors of Ref. [15] conclude that their experiments are consistent with the existence of the zero-temperature quantum phase transition .
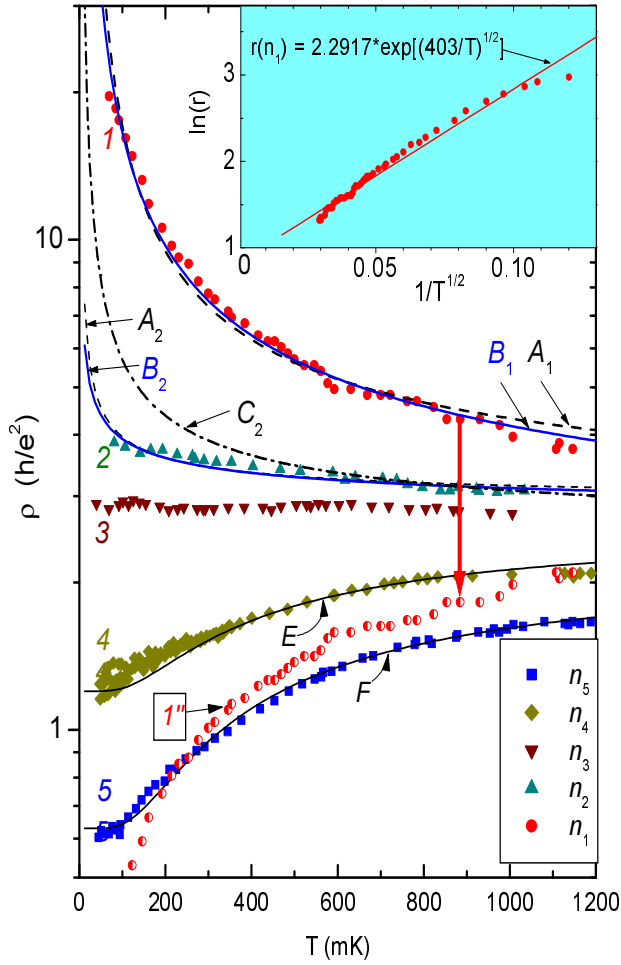
FIG. 1. Resistivity vs temperature, reproduced from Fig. 2 of Ref. [15]. Electron density n in units of $10^{10}\,\mathrm{cm}^{-2}$: *1* - 6.85, *2* - 7.17, *3* - 7.25, *4* - 7.57, *5* - 7.85. Dashed curves A_1 and A_2: fits of data *1* and *2* by $2.29\exp\left[(403/T)^{1/2}\right]$ and $2.84\exp\left[(11.11/T)^{1/2}\right]$, respectively ($T$ is in mK). Solid curves B_1 and B_2: fits of the same data by $1.25\exp\left[(1685/T)^{1/3}\right]$ and $2.55\exp\left[(8./T)^{1/3}\right]$, respectively. Semi-open symbols (1″): reflection of data points *1* about curve *3*. Curves E and F: best fits of data *4* and *5* by $\rho_4(T) = 1.2 + 1.45\exp(-450/T)$ and $\rho_5(T) = 0.63 + 1.6\exp(-485/T)$, respectively.

Note that a rather narrow range of densities was explored in Ref. [15]. To show the development of the MIT-phenomenon over a wider range of densities and temperatures, we present in Fig. 2 the results of Ref. [18] for a sample with a close value of μ_{peak}. The encircled region of the $\{T, n\}$ plane (region KK in this figure) indicates roughly the domain of parameters explored in Ref. [15]. The temperature in Fig. 2 is normalized to the Fermi energy E_F, in order to demonstrate that resistivities corresponding to quite different Fermi energies, and therefore r_s-values, exhibit similar T-dependences. A strong (50-100%) drop in ρ is still present for $n \lesssim 20 \times 10^{11}\,\mathrm{cm}^{-2}$, which corresponds to $r_s \sim 2$ and $\rho \sim 0.01h/e^2$. Thus,

contrary to the popular opinion, the resistance drop is not intrinsic only to the low density range (high r_s-values).
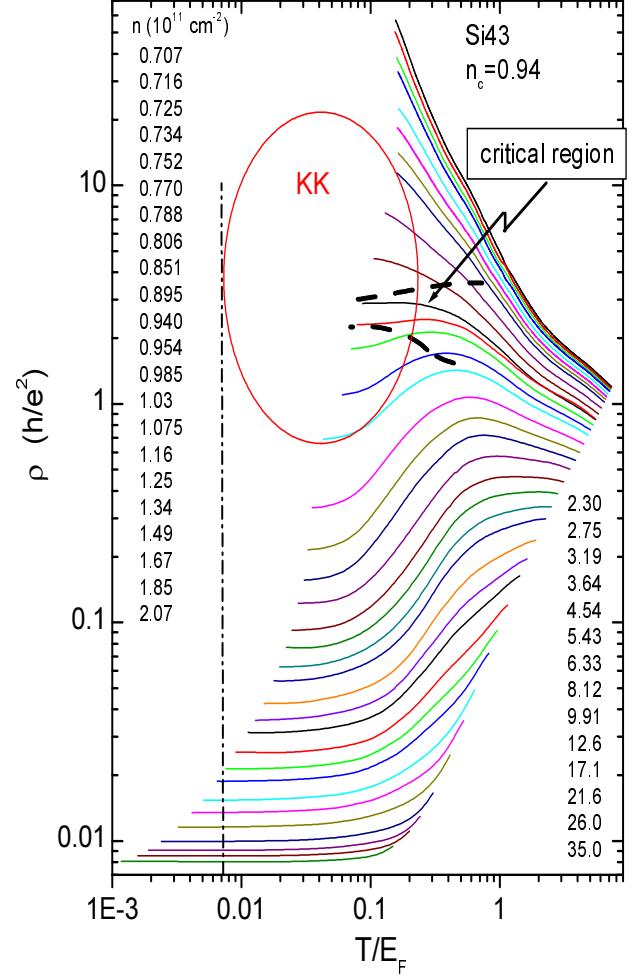

FIG. 2. Resistivity vs temperature (from 0.3 to $\approx 40\,\mathrm{K}$) for a wide range of densities and Fermi energies. The figure is reproduced from Ref. [18]. Encircled region: domain of parameters explored in Ref. [15]. Dash-dotted vertical line depicts the empirical temperature, $T_Q = 0.007E_F$, below which the logarithmic temperature dependence sets in.

We would like to stress once again that in the region of high densities and small resistivities quantum interference corrections of any kind (weak localization, exchange and Hartree interactions, spin-orbit, etc.) would amount only to a few percent variation (of both signs) in ρ with T. The phonon contribution to the resistivity of Si-MOSFETs can be shown to be negligible at least for $n \lesssim 20 \times 10^{11}\,\mathrm{cm}^{-2}$ in the relevant T-range. Therefore, there must be some other mechanism of the T-dependence at work, neither of quantum interference nor of phonon nature. The resistance drop of non-phonon nature persists up to the highest temperature of these measurements, which is $\approx 35 - 40\,\mathrm{K}$ for each curve in Fig. 2 (for conversion of densities into the Fermi-energies, use the formula: $E_F[\mathrm{K}] = 7.31 \times n[10^{11}\,\mathrm{cm}^{-2}]$).

3

II. METAL, INSULATOR, AND CRITICAL POINT. GENERAL DISCUSSION

Contrary to a common belief, a metallic phase and thus a metal-insulator transition can in principle occur in two dimensions. For example, in a hypothetical situation of non-interacting electrons spin-orbit coupling results in resistivity vanishing with temperature, provided that disorder is sufficiently weak ("weak antilocalization" [3,4,19]). However, if disorder is strong enough, it leads to localization. The metallic state is stable with respect to weak and short-ranged electron-electron interactions. The Coulomb interaction between electrons for small r_s overwhelms weak antilocalization and thus destroys the metallic state with zero residual resistivity. Nevertheless, the Coulomb interaction itself brings in an antilocalizing contribution to the resistivity (triplet channel). For small r_s, this contribution is smaller than the one from the singlet channel, and thus the net effect of interactions is the enhancement of localization. It may be the case that the triplet channel wins for larger r_s and a metallic state becomes possible. (In fact, this scenario follows from the RG treatment [20–22].) However, if such a state exists, it must be distinct from the conventional "disordered Fermi liquid".

So far, the experimental evidence is in contrast with the existence of a distinctly new state (see Refs.[5,7]). In fact, at higher densities the resistivity exhibits an insulating up-turn in agreement with the weak localization theory. We believe that in the existing experiments the true low-temperature asymptotic behavior has not yet been achieved for lower densities (in the vicinity of the transition). Therefore, one cannot interpret the crossover between $d\rho/dT > 0$ and $d\rho/dT < 0$ at finite temperatures as a proof for the existence of two distinct phases. Nevertheless, it is instructive to adopt such an interpretation for a moment and to compare the experimentally observed $\rho(T)$ with the commonly accepted picture of a quantum phase transition. Let us first recall what is known about the transition in 3D.

We start our analysis by formulating the definitions of metal and insulator. Arguing about definitions is not a too rational thing to do – a definition cannot be right or wrong. We simply point out that for the well-studied case of MIT in 3D, the commonly held definitions of both phases differ from those adopted by KK. In 3D systems, $\rho_{\text{insulator}}(T \to 0) \to \infty$, whereas $\rho_{\text{metal}}(T \to 0)$ is finite (see [3,4,22] and references therein). In fact, $d\rho/dT$ is negative in both the metallic and insulating phases, provided that the system is close to the transition point and the temperature is low enough.

For non-interacting particles, the conventional scenario of the 3D MIT is well supported by the perturbation theory and renormalization group arguments [3,4,23,22]. According to this scenario, in the metallic phase close to the transition $\rho(T)$ increases as $T^{-\alpha}$ as temperature decreases. The increase of $\rho(T)$ saturates at $T \lesssim T_{\text{sat}}$. Upon approaching the critical point, T_{sat}

tends to zero. Therefore, exactly at the critical density the $\rho(T)$-dependence is a power law: $\rho(T) \sim T^{-\alpha}$. In the insulating phase, the resistivity behaves as $\exp(T_0/T)$ (nearest neighbor hopping) or as $\exp[(T_0/T)^\beta]$ (variable range hopping). The exponent α is determined by mechanisms of dephasing. As a result, it is, strictly speaking, non-universal. Let the dephasing rate scale with temperature as $1/\tau_\varphi \propto T^p$. Then assuming that (a) one-parameter scaling holds, (b) this parameter is the conductance, and (c) $p = 1$, one arrives at $\alpha = 1/3$. Note that the electron-electron interactions can change the value of α.

Numerous experiments on 3D doped semiconductors confirmed the power-law behavior of ρ in the critical region. However, there is still no consensus regarding the value of α: both $\alpha = 1/2$ [24–26] and $\alpha = 1/3$ [27,28] have been reported. It is also possible that α depends on whether a semiconductor is compensated or not. (For the review of an MIT in 3D see Ref. [29].)

Returning to the MIT in 2D, we note that the assumption of a temperature-independent resistivity at the critical density is as doubtful as the statement that $\alpha = 1/3$ in 3D. Indeed, both of these points can be justified only within the one-parameter scaling. This scaling does not seem to apply universally even in the 3D case. For a 2D system of noninteracting electrons, this scaling predicts no metallic state at all. To get a chance to describe a metallic state, one needs to add more ingredients, e.g., electron-electron interactions, to the theory. Once we deal with a problem which is richer than localization in quenched disorder, there are no reasons to assume that the one-parameter scaling still holds (see, e.g., Ref. [20]). Therefore, a temperature-independent critical resistivity is an assumption rather than a law of nature.

As an example, consider the following density- and temperature-dependences of the resistivity (which do not follow from any of the existing theories, but do not contradict to any of the commonly respected general principles either):

$$\rho(n,T) = \rho_0(n) + \rho_1(n) \left[\frac{T}{T_1} \right]^{-\alpha} \exp \left[\frac{T_0(n)}{T} \right]. \quad (1)$$

It is natural to define the critical concentration, n_c, as the concentration at which $T_0(n)$ changes sign: $T_0(n_c) = 0$; for $n > n_c$ (metal), $T_0(n)$ is negative, whereas for $n > n_c$ (insulator), $T_0(n)$ is positive. It is also assumed that $\rho_0(n)$, $\rho_1(n)$ and $T_1(n)$ are smooth functions of the density in the vicinity of the transition point $n = n_c$. According to Eq.(1), exactly at the critical point $(n = n_c)$

$$\rho(n_c,T) = \rho_0 + \rho_1 \left(\frac{T}{T_1} \right)^{-\alpha}. \quad (2)$$

As we see, the critical resistivity is temperature-dependent and can even diverge as $T \to 0$. At the same time, the resistivity saturates at $\rho = \rho_0(n)$ for $T < -T_0$ in the metallic regime. (In our example, $\rho(T)$ has a

maximum at $T = -T_0/\alpha$. We do not think that such a maximum is a mandatory feature of the MIT in two dimensions.) We conclude that, at least in this example, a temperature-independent resistivity is a signature of a metal rather than of a critical point.

Regardless of this particular and rather artificial example, we notice that it is neither easy nor a straightforward task to determine the critical point of an MIT. Quantum phase transitions are zero temperature phenomena, whereas experiments are performed at finite T. Therefore, to determine the critical concentration, one does not have another choice but to analyze the data taken at the lowest accessible temperatures. This analysis should prove that $\rho(T)$ indeed acquires *insulating exponential behavior* as soon as the concentration gets lower than the apparent value of n_c.

We find it more appropriate to use the onset of the insulating exponential behavior in $\rho(T)$ rather than the vanishing of the derivative $d\rho/dT$ for an experimental identification of the critical point. We are not aware of any reasons to assign the meaning of the critical point to a density n_c^* at which the temperature dependence of the resistivity is least pronounced, as it was done by KK. Indeed, direct measurements of the two quantities, n_c^* (from the temperature independent "separatrix") and n_c (from the disappearance of the non-Ohmic hopping behavior) confirm their systematic difference for high-μ samples: n_c^* is larger than n_c by 1-5% (see, e.g., Fig. 3 in Ref. [30]). A similar difference arises also from the comparison of critical behaviors in the thermoelectric power and conductance [31]. Note that four out of five electron densities in Ref. [15], $n = n_1 - n_4$, fall into this ambiguous interval.

III. EXPERIMENTAL DETERMINATION OF THE CRITICAL POINT.

We now turn to the experimental data of Ref. [15]. The authors assume that n_c equals to 7.25×10^{10}cm^{-2} (curve 3 in Fig. 1). According to the definition of the critical point, proposed in Sec. II, this assumption implies that at lower densities ($n = n_2 = 7.17 \times 10^{10}$cm^{-2} and $n = n_1 = 6.85 \times 10^{10}$cm^{-2}), the resistivity increases exponentially as temperature decreases. It turns out that one can fit neither $\rho(T, n = n_2)$ nor $\rho(T, n = n_1)$ with a simple exponential dependence. However, $\rho(T, n = n_1)$ can be satisfactorily approximated by

$$\rho_{\exp}(T, n) = \rho_m^* \exp\left[(T_{0m}(n)/T)^m\right] \quad (3)$$

with either $m = 1/2$ or $m = 1/3$ (dashed line A_1 and solid line B_1 in Fig. 1, respectively). The quality of the fit can be seen from the deviations of the experimental data from the straight line in the inset to Fig. 1. Although curve 2 in Fiq. 1 does not seem to behave exponentially, one can still try to fit it by function (3) in two different ways:

i) one can assume that T_{0m} is proportional to $|n-n_c|$. (This dependence was observed in previous studies

of the MIT in Si MOSFETs [32–34]). Then, using the value of T_{0m} for $n = n_1$, one determines T_{0p} for $n = n_2$ by linear extrapolation, whereas ρ_m^* is found by fitting the experimental data. The results of this procedure for $m = 1/2$ are shown in Fig. 1 (dashed-and-dotted line C_2). It is quite clear that such a fit does not work (the same is true for $m = 1/3$, not shown).

ii) Alternatively, one can simply make the best fit of curve 2 by function (3) for $m = 1/2, 1/3$, treating T_{0m} and ρ_m^* as independent parameters. Results of this fit are depicted in Fig. 1 by curves A_2 and B_2, respectively. It is customary to assume that in the vicinity of the critical point, T_{0m} scales as $|n - n_c|^{z\nu}$. Given the values of T_{0m} for $n = n_{1,2}$, one can follow the assumption of KK that $n_c = n_3$ and find

$$z\nu = \frac{[\ln T_{0m}(n_1)/T_{0m}(n_2)]}{\ln\left[(n_3 - n_1)/(n_3 - n_2)\right]} = \quad (4)$$
$$= \begin{cases} 2.23 & \text{for} \quad m = 1/2; \\ 3.57 & \text{for} \quad m = 1/3. \end{cases}$$

It is quite alarming that these values of the critical exponent are substantially different from $z\nu = 1.2 - 1.6$ found in previous studies [33,34].

It is also often assumed that in the vicinity of the critical point $\rho(T, n)$ should be symmetric with respect to reflection about the separatrix [34–36]. To check if this relation works, we reflect curve 1 ($|\delta n| = n_3 - n_2 = 0.4 \times 10^{10}$ cm^{-2}) about the would-be separatrix (curve 3). The reflected curve is shown by semi-open circles. Had reflection symmetry worked, the reflected curve must have been located in between curves 4 ($\delta n = 0.32$ cm^{-2}) and 5 ($\delta n = 0.6$ cm^{-2}), closer to curve 4. In fact, the reflected curve *crosses* both curves 4 and 5. Thus by this criterion n_3 is not a critical density.

Unfortunately, for $n = n_2$ KK present the data only down to 80 mK, whereas for the rest of the densities, both smaller (n_1) and larger (n_3, n_4, n_5) than n_2, is shown down to 35 mK. If observed, a substantial increase in $\rho(T, n_2)$ with cooling from 80 mK to 35 mK would provide KK with a strong argument in favor of n_3 being close to the critical density.

Within the assumption that the data presented in Fig. 1 can at all be interpreted within a paradigm of a quantum phase transition, the contradictions demonstrated above suggest at least one of the following three conclusions:

i) the separatrix corresponds not to density n_3 but to some smaller density;

ii) the critical exponent is not universal for a given material but depends on the sample preparation, geometry, etc.;

iii) the exponential increase of the resistivity at density n_2 is suppressed due to overheating (we will discuss this option in Sec. V in more details).

Alternatively, one may conclude that the data of Ref. [15] do not provide evidence for a quantum phase transition between two distinct states of matter, but rather indicate a crossover between two different types of (relatively) high temperature behavior.

IV. RESISTIVITY OF SI MOSFETS IN THE CRITICAL REGION

As it has already been mentioned, any scenario of a quantum phase transition implies that for the densities close to the critical point n_c the resistivity $\rho(T, n)$ at sufficiently high temperatures demonstrates a critical behavior. This means that the deviation of $\rho(T, n)$ from the separatrix $\rho(T, n_c)$ is small for small values of the critical parameter $|n - n_c|^{z\nu}/T$. Keeping this in mind, we now recall the main results of earlier experiments on the resistivity of high mobility Si MOSFETs in the critical region and compare them with KK's results.

Consider, for example, the resistivity of a sample with $\mu_{\text{peak}} = 55,000\,\text{cm}^2/\text{V s}$, reproduced from Ref. [30,37] in Fig. 3. Assume for a moment that metallic and insulating behaviors can indeed be distinguished by the sign of the derivative $d\rho/dT$, as suggested by KK. At first glance, the application of this criterion to the region $T < 1\,\text{K}$ in Fig. 3 is rather unambiguous. Indeed, n_c^* determined from the condition of T-independent ρ corresponds to the dashed separatrix.
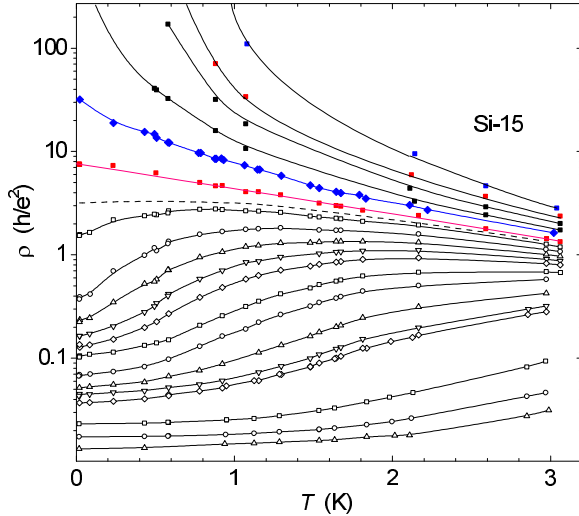

FIG. 3. $\rho(T)$ over the range from 0.016 to 3K. Different symbols correspond to n from 0.449 to 0.989 (in steps of 0.054), and further 1.1, 1.2, 1.42, 1.64, 1.74, 2.82, 3.9, 4.98 $\times 10^{11}\,\text{cm}^{-2}$. Reproduced from Refs. [30,37].

On the other hand, using the onset of activated and non-Ohmic conduction as a criterion for the transition, we find that the critical density n_c is lower, $n \approx 0.719 \times 10^{11}\,\text{cm}^{-2}$ and corresponds to the "tilted" (6th) curve from the top. At higher temperatures, some of the "metallic" curves change their slopes and follow quite closely those "insulating" curves, which correspond to lower densities (see solid curves in Fig. 3.) The region of

the $\{T, n\}$ plane, where such a behavior occurs, is defined as the critical region. Certainly, the density range of the critical region is narrower for lower temperatures.

As a rule, *the resistivity depends on the temperature in this region.* This feature, which can be found in almost any published data, is demonstrated in Figs. 2 and 3. For example, $\rho(T, n)$ in Fig. 3 increases by approximately a factor of 2, as T changes from 3K to 1K, both for $n = 0.719 \times 10^{11}\,\text{cm}^{-2}$ (the 6th curve from the top) and $n = 0.773 \times 10^{11}\,\text{cm}^{-2}$ (the 7th curve).

In fact, the $\rho(T)$-dependence at higher temperatures is usually more pronounced in higher-mobility samples. This is illustrated by Fig. 4, which is reproduced from Ref. [32].
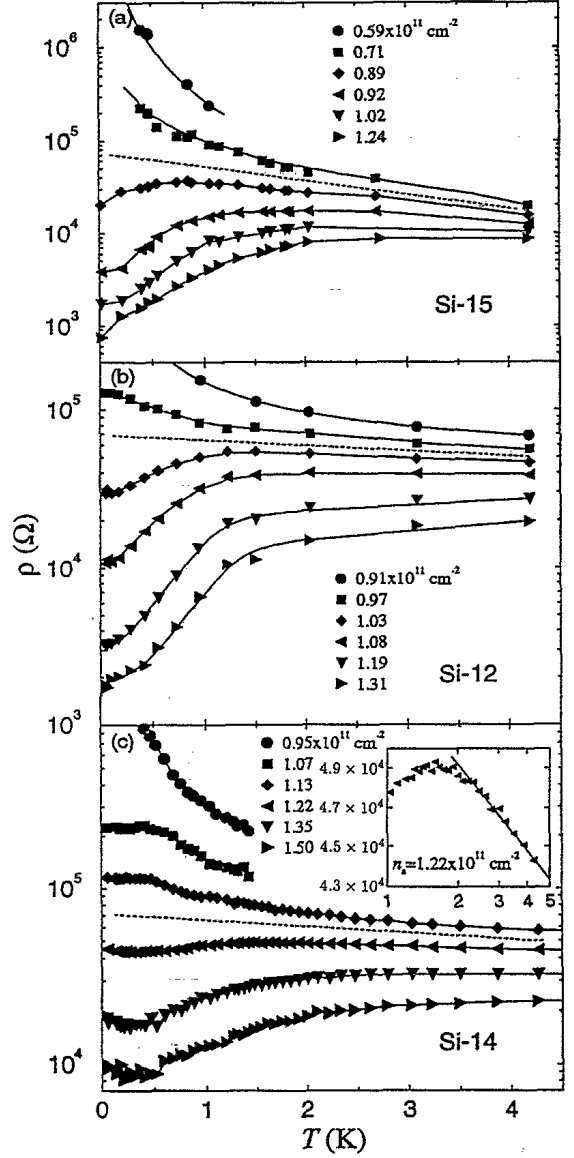

FIG. 4. Resistivity vs T near n_c for 3 samples with mobility decreasing from top to bottom. Reproduced from Ref. [32].

The slope of the dashed line for the sample with $\mu_{\mathrm{peak}} = 71,000\,\mathrm{cm^2/Vs}$ (Fig. 4a) is negative and its absolute value is about 10 times bigger than in Fig. 4b ($\mu_{\mathrm{peak}} = 33,000\,\mathrm{cm^2/Vs}$). The reduction of the slope with μ_{peak} persists down to relatively low mobilities until eventually the slope changes sign. This is seen from Fig. 5a, reproduced from Ref. [37], where the authors presented $\rho(T)$ for a sample with μ_{peak} as small as $5,000$ cm^2/Vs.
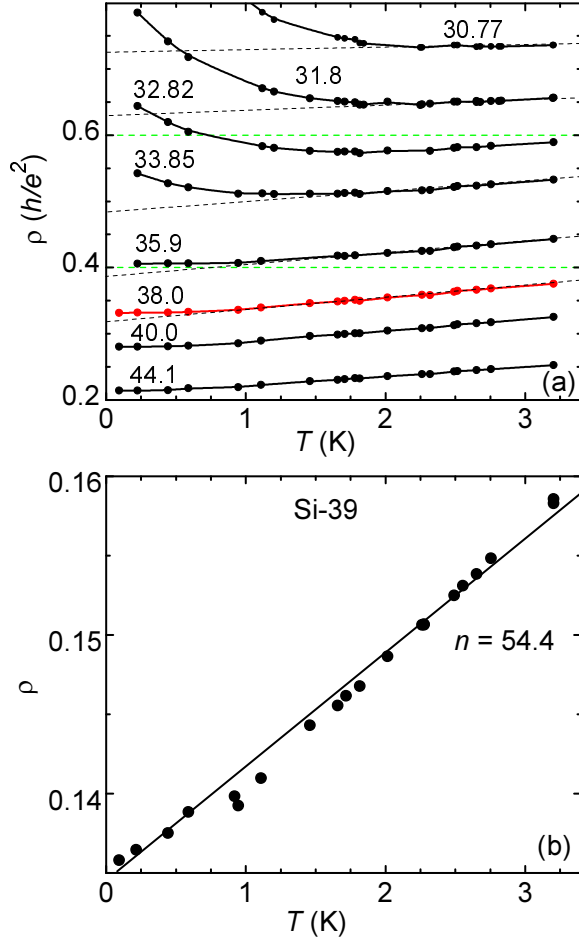

FIG. 5. Resistivity vs temperature for a *low mobility sample* (Si-39). Numbers at the curve indicate the densities in $10^{10}\mathrm{cm^{-2}}$. Reproduced from Ref. [37].

Even for this low-mobility sample, the $\rho(T)$-dependence at the highest density measured ($n = 54.4 \times 10^{11}\mathrm{cm^{-2}}$) shows no signs of saturation in the accessible temperature region (above $100\,\mathrm{mK}$), see Fig. 5b. There is no consistent theoretical understanding of this μ_{peak}-dependence of the slope yet. Such an understanding will probably come together with the description of the "critical resistance", which apparently changes from about $0.3e^2/h$ at low mobilities to several units of e^2/h at high μ_{peak}[30,37], (see also discussion in Refs. [38,39]). On the other hand, it is natural that *for the intermediate*

mobilities, the slope is small, because it should vanish somewhere in the range from $5,000$ to $30,000\,\mathrm{cm^2/Vs}$. The mobility of KK's sample is $27,000\,\mathrm{cm^2/Vs}$, i.e., right in this range. Therefore one should not be too surprised by the fact that *in some range of the concentrations the temperature dependence of the resistivity is rather flat*, as it was observed by KK.

Last but not least: a variety of the $\rho(T)$- dependences in the critical region indicates by itself the absence of universality in observed MITs [30] and raises doubts that Si MOSFETs undergo a genuine quantum phase transition.

V. "HOW COLD ARE THE ELECTRONS?"

Heating of electrons by the applied source-drain field as well as by external noise is a common problem in low-temperature transport measurements. This problem turns out to be especially serious in Si MOSFETS in the vicinity of the "metal-insulator transition" and at temperatures $\lesssim 100\,\mathrm{mK}$ due to *) weak temperature dependence of the resistivity in this range of densities and **) weak electron-phonon coupling. As a result, the interpretation of resistivity measurements becomes rather ambiguous. For example, the temperature interval, in which the resistivity appears to be temperature-independent, may look to be much broader than it actually is. The goal of this section is to demonstrate the seriousness of this problem.

In macroscopic Si MOSFETs, as well as in bulk Si, electrons couple to phonons only via the deformation potential. As this coupling is rather weak, the electron temperature, T_e, may exceed substantially the bath temperature, T_{bath}; it is also not easy to control T_e. To provide convincing data, one has to use a "thermometer" capable of measuring T_e directly. Any observable which depends on T_e, other than the (zero-magnetic-field) resistivity itself, can be used for this purpose. In earlier studies on high mobility Si-MOSFETs, Refs. [30,37,40–42], which reported results of the resistivity measurements down to $T_{\mathrm{bath}} = 18 - 25\,\mathrm{mK}$, the electron temperature was attempted to be measured via
i) the amplitude of Shubnikov - de Haas (ShdH) oscillations,
ii) temperature dependence of the hopping resistivity,
iii) linearity of $I - V$-characteristics, and
iv) T_e-dependence of the phase relaxation time.
Unfortunately, results on T_e in the up-to-date measurements of 2D MIT are neither fully consistent with each other nor convincing enough below $T_{\mathrm{bath}} = 250\,\mathrm{mK}$.

No independent measurement of T_e is reported in Ref. [15]. It is only mentioned that the source-to-drain bias, U, was less than $200\,\mu\mathrm{V}$ "to ensure that the total power dissipated in the sample was less than $10^{-13}\,\mathrm{W}$". As it is not clear whether this is enough to prevent electrons from being overheated, it is worth discussing electron heating in Si MOSFETs.

Electrons are driven out of the equilibrium by the applied voltage and/or by external noise. Let P be the power deposited in the electron system. In the stationary state, all this power leaves the system either with electrons (through contacts) or with phonons. The phonon mechanism dominates at higher temperatures. If T is low enough, this mechanism can be neglected compared to electron out-diffusion, in which thermal conductivity of the electrons determines the heat balance. As our task is to estimate heating at $T \sim 100$ mK, we first discuss what happens without phonons.

Given the temperature increment $\Delta T = T_e - T_{\text{bath}}$, the power, which is carried out by electrons through the leads, can be estimated as (see Refs. [43] and also [44] for recent discussion):

$$P = \left(\frac{2\pi}{e}\right)^2 \frac{T_{\text{bath}}\Delta T}{R}, \qquad (5)$$

where R is the resistance which differs from the resistivity by the aspect ratio. To obtain a lower bound estimate of the bath temperature T_{h} at which heating becomes strong, i.e., $\Delta T \simeq T_{\text{h}}$, we neglect heating due to external noise and assume that the main reason for heating is the source-to-drain bias, U. Accordingly, $P = U^2/R$. Expressing the ratio $\Delta T/T_{\text{bath}}$ through P via Eq. (5), we obtain

$$\frac{\Delta T}{T_{\text{bath}}} = PR\left(\frac{e}{2\pi T_{\text{bath}}}\right)^2 = \left(\frac{eU}{2\pi T_{\text{bath}}}\right)^2. \qquad (6)$$

Strictly speaking, Eqs.(5,6) are valid only when $\Delta T \ll T_{\text{bath}}$. Nevertheless, one can use them to estimate T_{h} as

$$T_{\text{h}}[\text{mK}] = 1.8U[\mu\text{V}]. \qquad (7)$$

For $U \sim 200\mu$V, Eq. (7) gives $T_{\text{h}} = 360$ mK. For lower bias, e.g., for $U = 50\,\mu$V this estimate gives $T_{\text{h}} = 90$ mK, whereas the experimental dependence $T_e(P)$ [45] taken at $T_{\text{bath}} = 140$ mK shows that at this bias electrons are already overheated: $T_e = 1.4T_{\text{bath}}$. Despite its simple form, Eq. (7) is quite general and universal. For example, it is also valid in the presence of Fermi-liquid interactions, which do not affect the Wiedemann-Franz law [46].

Taking into account the phonon mechanism of energy relaxation does not change our estimate of characteristic temperature T_{h} substantially. This is so because T_{h} is low enough for out-diffusion to dominate over phonon emission. Indeed, the electron-phonon energy loss rate for weak overheating can be written as

$$P_{\text{ph}} = \frac{\pi^2}{3}\nu_F LW \frac{T_{\text{bath}}\Delta T}{\tau_{\text{e-ph}}}, \qquad (8)$$

where ν_F is the density of states at the Fermi level, L and W are the sample length and width, respectively, and $\tau_{\text{e-ph}}$ is the electron-phonon relaxation time. We assume that electrons are coupled to phonons via the screened deformation potential. The corresponding relaxation time for $T_{\text{bath}} \ll \hbar k_F s = 3.65\sqrt{n}$ [K] (where s is the speed of sound, and n in 10^{11}cm^{-2}) can be written as [47,48]

$$\frac{1}{\tau_{\text{e-ph}}} = \frac{120}{\pi}\frac{\Xi^2 m^* k_F^2 a_0^3}{\hbar^2 M s}\left(\frac{T_{\text{bath}}}{k_F s}\right)^3\left(\frac{T_{\text{bath}}}{\kappa s}\right)^2, \qquad (9)$$

where Ξ is the deformation potential constant, M is the atomic mass, a_0 is the effective lattice constant defined in such a way that M/a_0^3 is equal to the mass density, and κ is the inverse screening length. Comparing (5) to (8), one finds that out-diffusion dominates, provided that T_{bath} is smaller than

$$T^* = \frac{\pi}{5}\hbar k_F s\frac{h}{\rho e^2}\mathcal{E}\mathcal{P}r_s^2\frac{\Xi}{\varepsilon_F}\left(\frac{a_0}{L}\right)^2, \qquad (10)$$

where $\mathcal{E} \equiv Ms^2\left(\hbar^2/m^*a_0^2\right)^2/\Xi^3$ and $\mathcal{P} \equiv \hbar/m^*sa_0$. Substituting material constants for a Si MOSFET into (10), one obtains

$$T^*[\text{K}] = 0.3\left(\rho n^2\right)^{-1/5}L^{-2/5}, \qquad (11)$$

where ρ, n, and L are measured in h/e^2, 10^{11} cm^{-2}, and mm, respectively. Using experimental results on electron heating in Si MOSFETs [45], one estimates the observed crossover temperature as 0.7 K, whereas Eq. (11) gives $T^* = 0.6$ K for the experimental conditions of Ref. [45]. For the conditions of another experiment on heating [49], Eq. (11) gives $T^* = 0.3$ K, which is just the lowest temperature of this measurement. No clear crossover from the phonon to out-diffusion mechanisms was observed in Ref. [49] for $T \geq 0.3$ K (although the T-dependence of the energy-loss rate does start to slow down at $T \simeq 1$ K). This is again consistent with our estimate. We thus conclude that Eq. (11) is consistent with the experimental data and can serve as at least a lower bound for T^*.

For typical values of $\rho \simeq e^2/h$, $n \simeq 10^{11}$ cm^{-2}, and $L = 0.1 - 1$ mm, Eq. (11) gives $T^* \simeq 0.3 - 0.75$ K, which is of the same order as T_{h} estimated above as 0.36 K. Note that in deriving (11) we assumed that ρ was T-independent. Taking the metallic-like $\rho(T)$-dependence into account enhances out-diffusion at low temperatures and thus shifts T^* towards even higher values. Also, taking into account possible external noise would further increase the value of T_{h}. The lowest bath temperature in Ref. [15] is $T_{\text{b}} = 35$ mK. Using Eq. (7), we see that in order to prevent electron heating, the bias voltage has to be much smaller than 20 μV, which is ten times smaller than the upper boundary for U indicated in Ref. [15].

We thus see that there are serious reasons to doubt that the electron temperature in KK's measurements was below ~ 100 mK.

VI. TEMPERATURE-INDEPENDENT RESISTIVITY

Now it is time to analyze the first of the two main arguments that KK brought in favor of the "true quantum phase transition", namely, the temperature-independent resistivity at what they believe to be the critical point. More specifically, KK claim that they observed no T-dependence (within a 10% margin) in the interval $35\,\mathrm{mK} < T < 1\,\mathrm{K}$. They analyzed the data in terms of our recent theory of Anderson localization by temperature-dependent disorder [5] and came to the conclusion that within this theory a nearly constant $\rho(T)$ would imply a "remarkable coincidence".

We begin our discussion of this issue with summarizing briefly the argument of KK. Afterwards, we present our theoretical counterarguments.

The theory of Ref. [5] describes the T-dependence of the observable resistivity, ρ, in the presence of two factors: i) the T-dependence of the Drude resistivity, ρ_{d}, and ii) Anderson localization. Each of these two factors brings a T-dependence of its own and the resultant $\rho(T)$-dependence is described by the following scaling equation

$$\left[\frac{1}{\beta(\rho)} - \gamma\right] \frac{d\ln\rho}{d\ln T} = \frac{p}{2} + \left[\frac{1}{\beta(\rho_{\mathrm{d}})} - 1\right] \frac{d\ln\rho_{\mathrm{d}}}{d\ln T}, \quad (12)$$

where ρ is measured in units of h/e^2, $\beta(x)$ is the Gell-Mann–Low function, whereas p and γ parameterize the phase-breaking time as

$$\tau_\varphi \propto T^{-p}\rho^{1-2\gamma}. \quad (13)$$

The T-dependence of ρ_{d}, entering the RHS of Eq. (12), leads to a variety of behaviors in $\rho(T)$. In Ref. [5] we discussed in particular the situation when $\rho(T)$ exhibits a maximum at a rather high temperature $T_{\max}$ (close to the starting temperature of the RG flow).

Is a very slow variation of $\rho(T)$ over a wide T-interval permitted in this model? KK answer this question negatively. Their argument goes as follows: to have $d\rho/dT = 0$ within some interval of temperature, one has to require that the RHS of Eq. (12) is equal to zero within this interval. This is possible only if the Drude conductivity ρ_{d} is related in a specific way both to the β-function and to τ_φ. Such an exact relation is unlikely, given the different origin of these quantities.

This argument sounds to be formally correct. Nevertheless, as we will demonstrate shortly, it is possible to achieve a very small variation of $\rho(T)$ within a wide temperature interval without imposing any constraints of this kind on ρ_{d}. It turns out that the variation of $\rho(T)$ is small enough and the T-interval is wide enough to agree with the experiment. This possibility results from the fact that there is a whole family of the $\rho_{\mathrm{d}}(T)$-functions, which are parameterized by the electron density and some other parameters, such as the peak mobility. It is the freedom in fine-tuning these parameters

that allows one to suppress the variation of the observable resistivity dramatically.

For the purposes of this section, we adopt the following working definitions, consistent with those used in Ref. [15]: by "metal" ("insulator") we understand the region in which $d\rho/dT > 0(< 0)$ or $d\rho_{\mathrm{d}}/dT > 0$; by "critical region" we understand the region in which $\rho \simeq 1$.

As was discussed in Ref. [5], a maximum in $\rho(T)$ results from a competition between metallic-like $\rho_{\mathrm{d}}(T)$ and localization effects, the latter being controlled both by $\rho_{\mathrm{d}}(T)$ and τ_φ. At higher T, when ρ_{d} is sufficiently large, localization can already be strong enough to ensure the negative sign of $d\rho/dT$ despite $\rho_{\mathrm{d}}(T)$ decreasing with temperature. However, as T goes down, ρ_{d} decreases, localization weakens and $d\rho/dT$ changes its sign from negative to positive, thus a maximum in $\rho(T)$ occurs. This can happen provided that certain conditions are satisfied. In particular, one has to require that

$$\rho_{\mathrm{d}}(T_0) > \rho_c, \quad (14)$$

where T_0 is the temperature at which τ_φ is comparable to the transport mean free time (the flow of Eq. (12) starts at $T = T_0$) and ρ_c is some critical resistivity determined by a particular form of $\rho_{\mathrm{d}}(T)$. For example, if $\rho_{\mathrm{d}} \propto T^q$, critical resistivity ρ_c is a certain function of p/q (see Eq. (9) of Ref. [5]). In this case, there is another condition for the maximum in $\rho(T)$ to occur, namely, $p > 2q$.

As the resistivity $\rho_{\mathrm{d}}(T_0)$ depends on the electron density n, it can be tuned by the gate voltage. This tuning can drive the system between the domains of parameters corresponding either to a maximum or no maximum in $\rho(T)$. Let us start at $\rho_{\mathrm{d}}(T_0) > \rho_c$, so that the maximum in ρ is at $T = T_{\max}$. As ρ_{d} approaches ρ_c from above, $T_{\max}$ remains finite, while the maximum flattens out and disappears as soon as $\rho_{\mathrm{d}}(T_0)$ approaches ρ_c. At this moment, both $d\rho/dT$ and $d^2\rho/dT^2$ vanish, and the variation of ρ around $T_{\max}$ takes place solely due to higher ($n \geq 3$) terms in the Taylor expansion of ρ.

How large is the temperature interval, in which this variation does not exceed some given value? Denoting,

$$t \equiv \ln T,$$
$$t_{\max} \equiv \ln T_{\max},$$
$$\beta_{\mathrm{d}} \equiv \beta[\rho_{\mathrm{d}}(t_{\max})],$$
$$y(t) \equiv \ln \rho_{\mathrm{d}}(t),$$
$$\beta'_{\mathrm{d}} \equiv d\beta(\rho)/d\ln\rho\Big|_{\rho=\rho_{\mathrm{d}}},$$
$$\beta''_{\mathrm{d}} \equiv d^2\beta(\rho)/d\ln\rho^2\Big|_{\rho=\rho_{\mathrm{d}}},$$

we obtain from Eq. (12)

$$\frac{d\rho}{dT}\Big|_{T=T_{\max}} = 0 \Rightarrow \dot{y}(t_{\max}) = \frac{p}{2}\frac{\beta_{\mathrm{d}}}{\beta_{\mathrm{d}} - 1} \quad (15\mathrm{a})$$

$$\frac{d^2\rho}{dT^2}\Big|_{T=T_{\max}} = 0 \Rightarrow \ddot{y}(t_{\max}) = -\frac{\beta'_{\mathrm{d}}}{\beta_{\mathrm{d}}(\beta_{\mathrm{d}} - 1)}[\dot{y}(t_{\max})]^2. \quad (15\mathrm{b})$$

The third (logarithmic) derivative of ρ is then found from Eq. (12) to be

$$\left.\frac{d^3 \ln \rho}{dt^3}\right|_{t=t_{\max}} = (1/A)\left[S - \frac{\beta_{\mathrm{d}}-1}{\beta_{\mathrm{d}}}\left.\frac{d^3 y}{dt^3}\right|_{t=t_{\max}}\right], \quad (16)$$

where

$$A \equiv 1/\beta(\rho) - \gamma \qquad (17)$$

and

$$S \equiv \frac{2\beta_{\mathrm{d}}'^2 - \beta_{\mathrm{d}}\beta_{\mathrm{d}}''}{\beta_{\mathrm{d}}^{\,3}}\,[\dot{y}(t_{\max})]^3 - \frac{3\beta_{\mathrm{d}}'}{\beta_{\mathrm{d}}^2}\ddot{y}(t_{\max})\dot{y}(t_{\max}) \quad (18\mathrm{a})$$

$$= \frac{p^3}{8(\beta_{\mathrm{d}}-1)^4}\left[\left(2\beta_{\mathrm{d}}'^2 - \beta_{\mathrm{d}}\beta_{\mathrm{d}}''\right)(\beta_{\mathrm{d}}-1) + 3\beta_{\mathrm{d}}'\right]. \quad (18\mathrm{b})$$

(In going from (18a) to (18b), we used (15a,15b)).

For a sufficiently narrow interval of t around $t_{\max}$, it suffices to keep only the lowest (cubic) term in the Taylor expansion of $\ln \rho(T)$ as a function of $\ln(T/T_{\max})$:

$$\ln\left(\frac{\rho(t)}{\rho(t_{\max})}\right) = (1/6A)\left[S - \frac{\beta_{\mathrm{d}}-1}{\beta_{\mathrm{d}}}\frac{d^3 \ln \rho_{\mathrm{d}}}{dt^3}\big|_{t=t_{\max}}\right] \times$$

$$\times (t - t_{\max})^3. \qquad (19)$$

As y is supposed to be a smooth function of t, one can estimate $d^n y/dt^n$ as $\bar{y}/\bar{t}^n$, where $\bar{y}$ is the typical value of y in the interval $\bar{t}$. In the "critical region", $\rho \simeq 1$, which means that $\bar{y} \simeq 1$ as well, and hence $\beta_{\mathrm{d}} \simeq \beta_{\mathrm{d}}' \simeq \beta_{\mathrm{d}}'' \simeq 1$. Thus the two terms in the square brackets in (19) are of the same order and (19) reduces to

$$\ln \rho(t) - \ln \rho(t_{\max}) \simeq (S/6A)(t - t_{\max})^3. \qquad (20)$$

Let us now estimate the number D of decades in temperature

$$T_{\mathrm{high}}/T_{\mathrm{low}} = 10^D, \qquad (21)$$

for which

$$|\ln \rho(t) - \ln \rho(t_{\max})| \leq 10\%. \qquad (22)$$

(KK found that their ρ is T-independent with this accuracy.) Using an interpolation formula $\beta(\rho) = -\ln(1 + 2\rho/\pi)$ [50] and choosing $\rho_{\mathrm{d}}(T_{\max}) = \pi/2$, we have

$$D \lesssim 2.3|A|^{1/3}/p. \qquad (23)$$

For $p = 1$ and $|A| = 1$, we get $D \lesssim 2.3$, i.e., the resistivity stays within a 10%-margin for more than *two decades* in T. (Note that a numerical coefficient of the order of unity, which we neglected in deriving (20), would enter (23) only under the cubic root, and is thus unimportant.) Fig. 6 demonstrates $\rho(T)$ calculated from Eq. (12) for ρ_{d} taken from the model of Ref. [51]:

$$\rho_{\mathrm{d}} = \rho_1 + \bar{\rho}\,(T/T_0)^q$$
$$\times \begin{cases} (1 + |\delta|T_0/cT)^q\,, & \text{for } \delta \leq 0,\text{ "insulator"}; \\ e^{-\delta T_0/T}, & \text{for } \delta > 0,\text{ "metal"}, \end{cases} \quad (24)$$

where $\delta = \delta_0(n - n_c)/n_c$ characterizes the distance from the critical point and $c \simeq 1$. The dimensionless coefficient δ_0 depends on details of the model (cf. Eq.(11) of Ref. [5]). For the purposes of the present paper, we view Eq. (24) as simply a phenomenological form for $\rho_{\mathrm{d}}(T)$, regardless of the model [51] it was originally derived from. This form is consistent with the experiment, if $\delta_0 \simeq 1$.

The value of $\delta = 0$ corresponds to the central line in Fig. 6, which on this scale is essentially constant for more than two decades. Zooming in (cf. the inset in Fig. 6), one sees that in fact $\rho(\delta = 0, T)$ stays within a 10%-margin over exactly *three decades* in T.
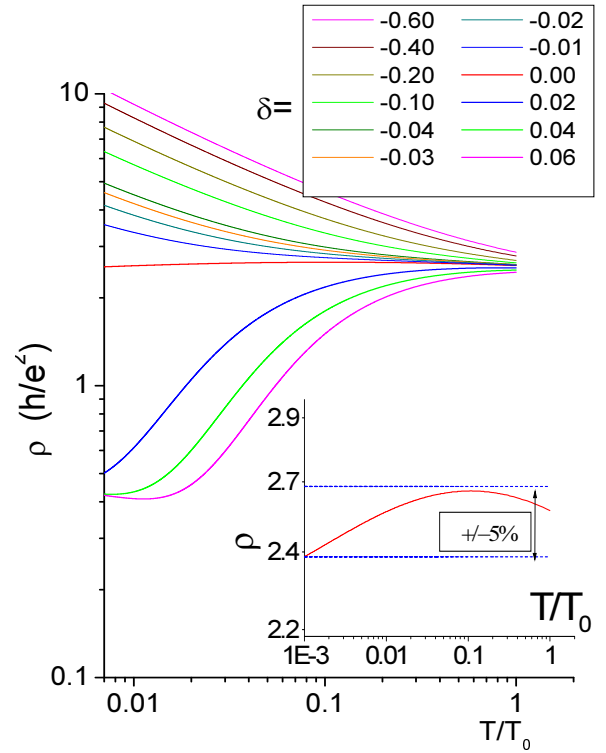

FIG. 6. Solution of Eq. (12) for different values of δ (distance from the "critical point"). $p = 1, q = 1/4, \gamma = 1, \rho_1 = 0.28, \bar{\rho} = 2.3$. Inset: the central (almost flat) curve ($\delta = 0$) over a wider temperature range.

KK claim that the observed $\rho(T)$ stays within this margin over only 1.5 decades, which is smaller than even the conservative estimate (23). (As it was discussed in Sec. V, a more realistic estimate of KK's T-interval is one decade). A decrease in δ by 1% ($\delta = -0.01$) leads to a weakly insulating behavior: $\rho(T)$ increases by 30% as T decreases by 1.5 decades. An increase in δ by 4% ($\delta = 0.04$) leads to a well-pronounced metallic behavior: $\rho(T)$ drops by about a factor of two over 1.5 decades in T. These features are in quantitative agreement with the experiment [15]. (For the sake of simplicity, we assumed that parameters ρ_1 and $\bar{\rho}$ do not depend on the density, i.e., on δ. Taking these dependences into account should lead to even better agreement with the experiment.)

We emphasize that *no fine tuning* between the β-function and ρ_d was used. If it is the competition between the decrease in Drude resistivity with temperature and the localization effects that leads to the maximum in the observable resistivity, then the $\rho(T)$-dependence flattens out, as the density approaches the threshold for a maximum. It turns out that if the exponent p is not too big (which is usually the case), this flattening suppresses the variation of $\rho(T)$ below the experimentally observable level over several decades in temperature. Even an oversimplified model [5] which does not take into account, e.g., electron-electron interactions (except for as a possible phase-breaking mechanism), can easily produce almost constant resistivity in a temperature interval, which is *two orders of magnitude* wider than the experimental one. Therefore, we do not find it too much of a "remarkable coincidence" that at some density the system demonstrates a fairly constant ρ.

We now turn to the data presented in Ref. [15]. The authors emphasize strongly that at $n = n_3 = 7.25 \times 10^{11}\,\mathrm{cm}^{-2}$ the resistivity is almost temperature independent in the temperature range $35\,\mathrm{mK}$ - $1\,\mathrm{K}$. They also add that $\rho(T)$ decreases with T for $T \gtrsim 1.8\,\mathrm{K}$. It would be interesting to know what happens at intermediate temperatures: $1\,\mathrm{K} < T < 1.8\,\mathrm{K}$. This point being unclear, KK fill in the missing temperature interval by combining their result with the one obtained on another system. Indeed, they write:

"In combination with the results of Ref. [10] (Ref. [52] of this paper–AMP) where the temperature-independent curve (with essentially the same value of resistivity) was observed in the temperature range 250 mk-1.8 K in another 2D system in silicon, we further allege that there is no observable T-dependence at $n_s = n_c$ in the temperature range 35 mk-1.8 K...". (Ref. [15], p. 3, second paragraph.)

Parenthetically, we note that the resistivities of combined curves differ by 35%. As far as we understand, temperature intervals of experiments on different systems are not additive parameters, and thus the procedure described above is not justified.

Summarizing this part of our discussion, we take the liberty to describe the experimental situation in the following way:

In a sample from the intermediate peak mobility range (where temperature dependence of ρ is known to be quite weak at high T), one can find a density such that ρ does not change for more than 10% within about one order of magnitude in T. This statement is a result of measurements on a single sample in the interval $100\,\mathrm{mK} < T < 1\,\mathrm{K}$ (when a realistic estimation of the electron heating is taken into account). (It will be two samples if results of Ref. [52] for $250\mathrm{mK} < T < 1.8\mathrm{K}$ are taken into consideration.)

In our opinion, there are neither theoretical nor experimental reasons to believe that the density corresponding to the weakest $\rho(T)$-dependence is indeed the critical one. i.e., that $\rho(T \to 0) \to \infty$ and $\rho(T \to 0)$ is finite for lower and higher densities, respectively. The lack of a pronounced temperature dependence of ρ in a limited range of T does not signal any remarkable phenomenon and is well-described by a simple theory of Anderson localization in a temperature-dependent disorder (Ref. [5]).

VII. ON THE APPARENT LINEAR TEMPERATURE DEPENDENCE OF $\rho(T)$

We now turn to point (2) of Ref. [15]] regarding the functional form of the observed $\rho(T)$-dependence. According to KK, this form is a) linear and b) qualitatively distinct from those observed by other authors. Fig. 7 reproduces Fig. 3b of Ref. [15], where the "almost linear dependence" is demonstrated. The data in Fig. 7 is data 5 of Fig. 1, displayed in truncated ($T < 400\,\mathrm{K}$) and full ($T < 1200\,\mathrm{mK}$) intervals. The truncated interval, in which $\rho(T)$ is supposed to be "almost linear", corresponds to Fig. 3b of Ref. [15].

We begin with an attempt to fit untruncated data 5 of Figs. 1, 7 in the whole interval $T < 1.2\,\mathrm{K}$ by an empirical expression [53]

$$\rho(T) = \rho_0 + \rho_1 \exp\left(-\frac{T_0}{T}\right), \qquad (25)$$

which is a simplified version of Eq. (1) with $\alpha = 0$. This expression is known to work reasonably well for the metallic phase not only in Si MOSFETs [30,53] but also in other 2D systems exhibiting the MIT-like behavior [35].

As is seen from Fig. 7, this attempt is quite successful. We achieved more than just acceptable fit in a much broader temperature interval than the one in which $\rho(T)$ looks "almost linear". Of course, it is not a much of an achievement to fit a smooth curve by a function with three free parameters (to demonstrate a flexibility in choosing the fitting parameters, Fig. 7 presents two sets of parameters that can be used). On the other hand, we see *no reason* to argue that the $\rho(T)$-dependence of Ref.

[15] at $n = n_4$ and $n = n_5$ (see Fig. 1) is qualitatively different from those observed in other experiments.
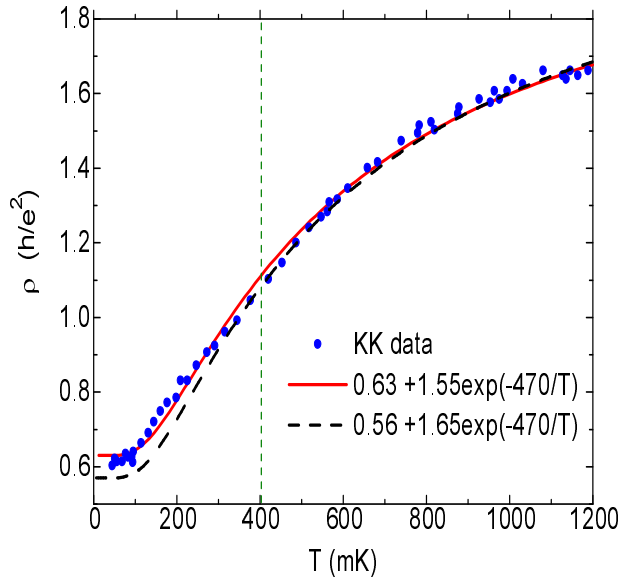

FIG. 7. Temperature dependence reproduced from Figs. 1, 3 of Ref. [15]. Region to the left of the vertical dashed line is the truncated T-range in Fig. 3a of Ref. [15], in which $\rho(T)$ is claimed to be linear. Bold dashed and continuous curves: fits of the data with Eq. (25).

Moreover, such "almost linear" $\rho(T)$-dependences are *typical for low-mobility samples in the "metallic" regime.* For instance, compare Fig. 7 with Fig. 5b, in which $\rho(T)$ for a low-mobility sample is plotted over a much wider temperature range (90 mK - 4 K), and for n about 50% higher than what can be called a critical concentration. The similarity of these two plots is quite clear.

We find it quite unlikely that the quasi-linear $\rho(T)$-dependences of Figs. 1 and 7 (as well as Fig. 5) manifest an exotic non-fermi-liquid state. Taking into account the whole set of features of the metallic state in Si-MOSFETs (which we listed in Ref. [5]), we conclude that these dependences must have a classical explanation (see the next paragraph for our definition of this term). For example, one cannot exclude that such behavior is caused by the temperature dependence of the screening length [55–58] or by the temperature dependence of the charge traps' population in the oxide [51].

We conclude this discussion by remarking that there seems to be some confusion in recent literature on MIT in 2D (see, e.g., Ref. [16]) regarding the distinction between "quantum" and "classical" behaviors. We believe that this confusion is mainly semantic. We call a regime "classical" as long as electron transport can be described in the framework of the Boltzmann theory. This does not necessarily mean that electrons obey Boltzmann (non-degenerate) statistics -in fact, the Fermi-liquid theory was originally formulated in terms of the Boltzmann equation [59]. It also does not exclude that a scattering cross-section is described in terms of quantum mechanics. (A more common name for this regime is "semi-classical" [60], although the semi-classical approximation may not necessarily be employed for the calculation of a scattering cross-section.) However, quantum-mechanical effects arising due to interference of electron waves, e.g., weak localization, cannot be described by the Boltzmann equation. We call the regime "quantum", if these effects are important. Although we agree with the authors of Ref. [16] in that "the formation of a Fermi surface" is an "intrinsic quantum-mechanical effect", this does not imply that the metallic-like temperature dependence of the resistivity, whose onset is correlated with the Fermi-surface formation, is of the quantum nature, in a sense of the definition given above.

VIII. AT WHAT TEMPERATURES WEAK LOCALIZATION MAY BE OBSERVED

There is one more point to discuss regarding the metallic behavior of Si MOSFETs. In addition to the almost linear dependence with $d\rho/dT > 0$, KK emphasize that no signs of weak localization (WL) was observed, although they went to temperatures as low as $T \sim 0.008 T_F$, where T_F is the Fermi temperature. Giving the authors the benefit of the doubt, we assume the most favorable scenario in which electrons are not overheated. Nevertheless, under this assumption the apparent absence of WL is not as surprising as KK present it. In fact, we believe that even the *bath* temperature was too high to observe the localization upturn in the T-dependence of the resistivity. Indeed, experimental studies of high-mobility Si MOSFETs in the wide range of electron densities [62,61] show that WL effects are well-pronounced in the (perpendicular) *magnetic field* dependence of the resistivity. The observed magnetoresistance is in a quantitative agreement with the conventional theory [3,4], even for densities rather close to the critical one. On the other hand, even for $n \gg n_c$ observation of the WL effect in the zero-field $\rho(T)$-dependence [42,8] (as well as the effects of electron-electron interactions) requires a precision of the order of $\delta\rho(T)/\rho(T) \ll (1/\pi)(e^2/h)\rho \sim 0.1\%$.

It turns out that the observed dependence in the range of "high temperatures", $T = (0.1 - 0.5)E_F$, is too strong to be explained by WL and quantum interaction effects, and can be attributed to the T-dependence of the Drude resistivity. However, the slope $|d\rho/dT|$ decreases with T. For relatively large densities, one can reach the temperature region, $T < T_Q$, in which the slope becomes comparable to that of the WL correction, $|d\rho/dT|_{WL} \sim (e^2/\pi)\rho^2 T$. In this region, quantum interference should contribute significantly to the $\rho(T)$-dependence. Authors of Ref. [8,42] found an empirical relation: $T_Q \lesssim 0.007 E_F$ (see Fig.(2)). As E_F is proportional to n, it becomes progressively more challenging to cool electrons below T_Q as n_c is approached. $E_F \simeq 5.5$ K for the highest electron

density reported by KK and thus $T_Q = 40$ mK. Therefore, in order just to reach the high temperature edge of the quantum transport region, where WL effects are seen, electrons should be cooled below 40 mK. The localization upturn may be expected to occur only at temperatures substantially smaller than even this one. In other words, the lowest temperature reported by KK (35 mK) is still too high for localization effects to be observed.

Why WL manifests itself so differently in the magnetoresistance and in the $\rho(T)$? This can be explained naturally by assuming that the Drude resistivity has a pronounced T-dependence, which masks quantum corrections. On the other hand, ρ_d is not expected to vary substantially in weak magnetic fields, thus WL can be seen in magnetoresistance.

IX. CONCLUSIONS

Some concluding remarks:

i) the apparent metal-insulator transition in 2D is a very interesting and unexpected phenomenon. Although we have limited our discussion to Si MOSFETs, which exhibit the strongest, among other 2D systems, "anomalous" metallic behavior, there is a whole variety of interesting effects not only in transport but also thermodynamic properties, e.g., compressibility [63,64]), observed in various 2D heterostructures. Whether all these observations have a universal explanation or there is a number of different, system-specific mechanisms at work, remains to be seen.

ii) we do not believe that the experimental results accumulated up-to-date provide a convincing evidence for this phenomenon being a true quantum phase transition between two distinct states of matter. More precisely, there is no evidence that the "metallic phase" (as defined by the sign of $d\rho/dT$) is a new state of matter. On the contrary, a whole set of features–see Refs.[5,7]–is consistent with a conventional behavior of a disordered Fermi liquid. This statement does not negate i). There is still no consensus on the origin of the effect, and this problem requires the most serious and intensive investigations.

iii) For a successful resolution of the problem, it is probably not sufficient to concentrate on the temperature dependence of the resistivity in the vicinity of the transition. Such features as a weak temperature dependence of the resistivity at a certain electron density in a limited (though large) temperature interval allow for different interpretations and, unfortunately, can not provide an unambiguous information about the zero-T state of the system. In particular, it is impossible to determine

even the critical density without making rather arbitrary assumptions.

iv) There is a serious experimental difficulty that up to now prevented a substantial increase of the temperature interval, in which the resistivity can reliably be measured in Si-MOSFETs. It turns out that in existing samples phonons can hardly cool the electron gas below a fraction of a Kelvin. On the other hand, the systems are quite noisy, and thus the applied voltage cannot be reduced much below the currently used level. As a result, even in the absence of non-equilibrium noise, one cannot neglect electron heating at $T \simeq 100$ mK.

[1] V. Senz, T. Heinzel, T. Ihn, and K. Ensslin, cond-mat/9910228.

[2] S. L. Sondhi, S. M. Girvin, J. P. Carini, D. Shahar, Rev. Mod. Phys. **69**, 315 (1997).

[3] P. A. Lee, T. V. Ramakrishnan, Rev. Mod. Phys. **57**, 287 (1985).

[4] B. L. Altshuler and A. G. Aronov, in *Electron-electron interactions in disordered systems*, ed. by A. L. Efros and M. Pollak (Elsevier, Amsterdam, 1985), p. 1.

[5] B. L. Altshuler, D. L. Maslov, V. M. Pudalov, cond-mat/9909353.

[6] We wish to emphasize that what we call "Drude resistivity" is not necessarily something well-known and described in textbooks. It is quite possible that the mechanism(s) of the unusual Drude resistivity in Si MOSFETs and other 2D systems is (are) based on new physics. The only one feature of the Drude resistivity which we retain in our argument is that ρ_d has nothing to do with quantum interference.

[7] Y. Meir, cond-mat/9912423.

[8] V. M. Pudalov, G. Brunthaler, A. Prinz, and G. Bauer, Phys. Rev. B **60**, R2154 (1999).

[9] V. M. Pudalov, G. Brunthaler, A. Prinz, G. Bauer, Pis'ma ZhETF **70**, 48 (1999) [JETP Lett. **70**, 48 (1999)].

[10] M. Y. Simmons, A. R. Hamilton, M. Pepper, E. H. Linfield, P. D. Rose, and D. A. Ritchie, cond-mat/9910368.

[11] T. M. Klapwijk and S. Das Sarma, Sol. St. Commun. **110**, 581 (1999).

[12] S. Das Sarma and E. H. Hwang, Phys. Rev. Lett. **83**, 164 (1999).

[13] Y. Meir, Phys. Rev. Lett. **83**, 3506 (1999).

[14] Y. Yaish, O. Prus, E. Buchstab, S. Shapira, G. Ben Yosef, U. Sivan, and A. Stern, cond-mat/9904324.

[15] S. V. Kravchenko and T. M. Klapwijk, cond-mat/9909458.

[16] P. Phillips, S. Sachdev, S. V. Kravchenko, and A. Yazdani, Proc. Natl. Acad. Sci. USA **96**, 9983 (1999) [cond-mat/9902025].

[17] S. V. Kravchenko, D. Simonian, K. Mertes, M. P. Sarachik and T. M. Klapwijk, Phys. Rev. B **59**, R12740 (1999).

[18] A. Prinz, V. M. Pudalov, G. Brunthaler, G. Bauer, in: Proc. Intern. Meeting SIMD-99, Mauri, (1999), edited by K.Hess. To be published in *Superlattices and Microstructures* (2000).

[19] S. A. Hikami, A. I. Larkin, and Y. Nagaoka, Prog. Theor. Phys. Lett. **63**, 707 (1980).

[20] A. M. Finkelstein, JETP **84**, 168 (1983). Z. Phys. B **56**, 189 (1984); Sov. Sci. Rev./Sec. A-Physics Reviews, edited by I. M. Khalatnikov, **14**, 3 (1990).

[21] C. Castellani, C. Di Castro, P. A. Lee, M. Ma, Phys. Rev. B **30**, 527 (1984); C. Castellani, G. Kotliar, P. A. Lee, Phys. Rev. Lett. **59**, 323 (1987). C. Castellani, C. Di Castro, H. Fukuyama, P. A. Lee, M. Ma, Phys. Rev. B **33**, 7277 (1986); C. Castellani, C. Di Castro, and P. A. Lee, *ibid.* **57**, R9381 (1998).

[22] D. Belitz, T. R. Kirkpatrick, Rev. Mod. Phys. **66**, 261 (1994).

[23] E. A. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, Phys. Rev. Lett. **42** 673 (1979). P. Lee and T. D. Ramakrishnan, Phys. Rev. B **26**, 4009 (1982).

[24] T. F. Rosenbaum, R. F. Milligan, M. A. Paalanen, G. A. Thomas, R. N. Bhatt, and W. Linn, Phys. Rev. **B27**, 7509 (1983).

[25] S. Hatsumoto, F. Komori, N. Sano, and S. Kobayashi, J. Phys. Soc. Jpn. , **56**, 2259 (1987).

[26] S. Bogdanovich, M. P. Sarachik, and R. N. Bhatt, Phys. Rev. B **60**, 2286 (1999).

[27] M. Maliepaard, M. Pepper, R. Newbury, J. E. F. Frost, D. C. Peacock, D. A. Ritchie, G. A. C. Jones, and G. Hill, Phys. Rev. **B39**, 1430 (1989).

[28] M. Watanabe, K. M. Itoh, Y. Outoka, and E. E. Haller, Phys. Rev. B **60**, 15817 (1999).

[29] M. P. Sarachik, in *Metal-Insulator Transition Revisited*, edited by P. P. Edwards and C. N. Rao (Taylor and Francis, London, 1995).

[30] V. M. Pudalov, G. Brunthaler, A. Prinz, and G. Bauer, JETP Lett., **68**, 442 (1998). [Pis'ma ZhETF **68**, 415 (1998)].

[31] R. Fletcher, A. D. B. Radcliffe, V. M. Pudalov, C. Possanzini, cond-mat/0002436.

[32] S. V. Kravchenko, G. V. Kravchenko, J. E. Furneaux, V. M. Pudalov, and M. D'Iorio, Phys. Rev. B **50**, 8039 (1994).

[33] S. V. Kravchenko, W. E. Mason, G. E. Bowker, J. E. Furneaux, V. M. Pudalov, and M. D'Iorio. Phys. Rev. B B **51**, 7038 (1995).

[34] S. V. Kravchenko, D. Simonian, M. P. Sarachik, W. Mason, and J. E. Furneaux, Phys. Rev. Lett. **77**, 4938 (1996).

[35] Y. Hanein, U. Meirav, D. Shahar, C. C. Li, D. C. Tsui, and H. Shtrikman, Phys. Rev. Lett. **80**, 1288 (1998).

[36] D. Simonian, S. V. Kravchenko, and M. P. Sarachik, Phys. Rev. B **55**, R13421 (1997).

[37] V. M. Pudalov, G. Brunthaler, A. Prinz, and G. Bauer, Physica E **3**, 79 (1998).

[38] S. V. Kravchenko, M. P. Sarachik, and D. Simonian, Phys. Rev. Lett. **83**, 2091.

[39] B. L. Altshuler and D. L. Maslov, Phys. Rev. Lett. **83**, 2092 (1999).

[40] V. M. Pudalov and S. T. Chui, Phys. Rev. B **49**, 14062 (1994); V. M. Pudalov, M. D'Iorio, S. V. Kravchenko, and J. W. Campbell, Phys. Rev. Lett. **70**, 1866 (1993).

[41] M. D'Iorio, V. M. Pudalov, and S. G. Semenchinsky, Phys. Rev. B **46**, 15992 (1992).

[42] V. M. Pudalov, G. Brunthaler, A. Prinz, and G. Bauer, JETP Lett., **68**, 534 (1998).

[43] D. E. Prober, Appl. Phys. Lett. **62**, 119 (1993); Phys. Rev. Lett. **75**, 3964 (1995).

[44] B. L. Altshuler, M. E. Gershenson, and I. L. Aleiner, Physica E, **3**, 58 (1998).

[45] B. J. Zieve, D. E. Prober, and R. G. Wheeler, Phys. Rev. B **57**, 2443 (1998).

[46] M. Reizer, Phys. Rev. B **44**, 12701 (1991).

[47] V. F. Gantmakher and Y. B. Levinson, *Carrier Scattering in Metals and Semiconductors* (North-Holland, Amsterdam), 1987.

[48] V. Karpus, Semicond. Sci. Technol. **5**, 691 (1990).

[49] R. Fletcher, V. M. Pudalov, Y. Feng, M. Tsaousidou, and P. N. Butcher, Phys. Rev. B **56**, 12422 (1997); Phys.Rev. B **60**, 8392 (1999).

[50] M. Batsch, I. Kh. Zharekeshev, and B. Kramer, Sol. State Commun.**95**, 353 (1995); I. Kh. Zharekeshev, private communication.

[51] B. L. Altshuler and D. L. Maslov, Phys. Rev. Lett. **82**, 145 (1999).

[52] M. P. Sarachik and S. V. Kravchenko, Proc. Natl. Acad. Sci. USA **96**, 5900 (1999) [cond-mat/9903292].

[53] V. M. Pudalov, JETP Lett. **66** 175 (1997). [Pis'ma ZhETF **66** 168 (1997)].

[54] A. P. Mills, Jr., A. P. Ramirez, L N. Pfeiffer, and K. W. West, Phys. Rev. Lett. **83**, p. 2805 (1999).

[55] A. Gold and V. T. Dolgopolov, Phys. Rev. B **33**, 1076 (1986)

[56] S. Das Sarma, Phys. Rev. B **33**, 5401 (1986).

[57] T. M. Klapwijk and S. Das Sarma, Sol. St. Commun. **110**, 581 (1999).

[58] S. Das Sarma and E. H. Hwang, Phys. Rev. Lett. **83**, 164 (1999).

[59] L. D. Landau, JETP **3**, 920 (1956).

[60] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart, and Winston, New York), 1976, Ch. 12.

14

[61] G. Brunthaler, A. Prinz, G. Bauer, V. M. Pudalov, E. M. Dizhur, J. Jaroszynski, P. Glod, and T. Dietl, Annalen der Phys. (Leipzig), **8**, 579 (1999); V. M. Pudalov, G. Brunthaler, A. Prinz, and G. Bauer, Physica B **249-251**, 697 (1998).

[62] V. M. Pudalov, G. Brunthaler, A. Prinz, and G. Bauer, JETP Lett. **65**, 932 (1997); Physica B, **249-251**, 697 (1998).

[63] S. C. Dultz, and H. W. Jiang, cond-mat/9909314.

[64] S. Ilani, A. Yacoby, D. Mahalu, and H. Shtrikman, cond-mat/9910116.